



Development and Validation Technological Pedagogical Content Knowledge (TPACK) Instrument for Teacher Mathematics in Elementary School

Titik Suryani; Wardani Rahayu; Ari Saptono

Department of Educational Research and Evaluation, Universitas Negeri Jakarta, Indonesia

<http://dx.doi.org/10.18415/ijmmu.v8i8.2951>

Abstract

Measurement of teacher competence through TPACK is very important to do. However, the majority of the instruments used to measure TPACK for elementary school teachers are not yet in line with the needs of the 21st Century. This study aims to develop a TPACK instrument for elementary school teachers in a valid and reliable Mathematic lesson. This study uses the quantitative research. The study sample involved 1490 teacher in elementary school, grade IV, V, and VI teachers from state and private elementary schools with A level of accreditation, spread across fourteen sub-districts in the Kabupaten Bogor. Data analysis techniques use Winsteps 4.0.1. So that, overall, this instrument can be used to measure TPACK of elementary school teachers in the content of mathematics lessons. This instrument can be used to map teacher competencies based on TPACK values. Furthermore, if there are some weaknesses in the 7 aspects of the TPACK, policy makers can determine the follow-up that needs to be determined to further improve the competence of elementary school teachers. The dimensions that stand out in this research are lesson content, teaching approach, technology, pedagogy, content, and mathematics.

Keywords: *TPACK; Rasch; Quantitative Research*

Introduction

The development of science and technology brings the industrial revolution 4.0. Klaus Schwab describes the industrial revolution as an era where individuals move using digital technology and reality in their activities and managing their lives (Xu et al., 2018). The term industry 4.0 refers to the era of the physical, biological, and digital dimensions forming a blend that is difficult to distinguish (Putrawangsa & Hasanah, 2018). Today, digital technology affects the education system (Hy et al., 2015). Technology has a positive impact on learning, one of which is in mathematics (Sintawati & Indriani, 2019). Technology, becomes one of the important elements in teaching Mathematics to improve students' understanding (Niess et al., 2009).

Mathematical knowledge that is integrated into technology must be instilled since students are in elementary school (Ling Koh et al., 2014). Fusarelli (2008) emphasized that teachers' pedagogical abilities and the use of information and communication technology as instructional tools are factors that help teachers and schools meet the challenges of preparing students to improve the skills needed in the 21st century. The results of this study are also supported by the results of research Tay, Lim, Koh (2012) and Hennessy, Ruthven and Brindley (2005) which states that the use of ICT will increase the

understanding of the content of knowledge and pedagogical abilities of teachers in learning mathematics (Ariani, 2015).

The Indonesian Ministry of National Education requires that every teacher in Indonesia must have standard competencies, which include; mastery of knowledge (content), technology, pedagogy, culture, humanity, nationality and civilization (Nasional et al., 2012). These requirements are in accordance with the objectives of 21st century education, namely education is directed to develop mathematics and natural sciences accompanied by social and human sciences (humanities) that are balanced with technology (Badan Standar Nasional Pendidikan, 2010).

In order to achieve this goal, the Indonesian Ministry of Education and Culture held a Teacher Competency Test (UKG). The results of the National UKG show that the competence of primary school teachers in Indonesia is below the average, with an average score of 42.06 and in 2013 with an average score of 42.5 at the national level (Das et al., 2009). Based on these data, education in Indonesia needs to improve, guru harus memiliki kesadaran mendedikasikan dirinya meningkatkan kemampuan four teacher competencies that adapt to technological developments (Richardo, 2016).

In principle, teacher knowledge of integrating technology in mathematics learning can be measured using the Technological Pedagogical Content Knowledge framework (TPACK) (Niess et al., 2009). TPACK itself originated from Shulman's idea of Pedagogical Content Knowledge (PCK), then Mishra and Koehler added a technological aspect to PCK (Koh et al., 2015). TPACK is a framework of knowledge needed by teachers to integrate technology in learning effectively (Koehler, 2006). The knowledges in question are: (1) Technological Knowledge (TK) which includes knowledge in utilizing technology, (2) Pedagogical Knowledge (PK) in managing students and learning in class, and (3) Content Knowledge (CK), knowledge of the material being studied or taught to students, and knowledge of the relationship between the first three knowledges in facilitating students to learn (Harris et al., 2009). Subsequently, the three main component bases of TPACK developed into seven components; Content Knowledge (CK), Pedagogical Knowledge (PK), Technological Knowledge (TK), Technological Pedagogical Knowledge (TPK), Technological Content Knowledge (TCK), Pedagogical Content Knowledge (PCK), and Technological Pedagogical Content Knowledge (TPACK)(Koehler et al., 2017).

Several studies related to measuring TPACK capabilities have been carried out, including research to measure the TPACK of prospective Mathematics teachers in learning (Pramuk, Ergun, Cakir, Yilmaz, & Ayas, 2013; Purwaningsih & Yuliati, 2015; Schmidt et al., 2009; Sholihah, Yuliati, & Wartono, 2016; Srisaswasdi, 2012; Sumarsono, Malik, & Hernandez, 2014). Other studies to measure the TPACK of teachers in secondary schools include (Gunawan et al., 2020; Gao et al., 2011; Listiawan & Baskoro, 2015; Patra & Guha, 2017; Rafi & Sabrina, 2019). Referring to these studies, the authors are interested in developing the TPACK instrument which aims to measure the TPACK of elementary school teachers in Mathematics, this is because there are not many studies that measure the TPACK of elementary school teachers, especially Mathematics.

This study is a study that modifies the TPACK instrument sourced from the seven previous TPACK components. The instrument uses a Likert scale questionnaire of 1 to 5. The Likert scale is a psychometric scale commonly used in questionnaires and the most widely used in survey research (Pilihan et al., 2015). Scarvia B. Anderson stated that the instrument is said to be valid if the instrument can measure what it is intended to measure (Caturiyati, 2013). To be valid, the instrument must have a scale (Perera et al., 2018). The purpose of this study was to obtain information about the TPACK ability of elementary school teachers in mathematics. Furthermore, TPACK research can be used as a reference in improving the quality of education (Jang & Tsai, 2012). The government can determine teacher professionalism development policies by referring to the TPACK measurement (Dessty, 2018).

Method

Research based on post-positivism paradigm with quantitative type using survey methodology, technical data collection with a Likert scale questionnaire method 1 to 5. The sample was selected by simple random sampling based on the consideration of schools that have A accreditation and use the 2013 Curriculum. Simple random sampling is a technique of taking samples from each population unit by providing equal opportunities and in practice it is done by lottery or random numbers (Triyono, 2003). The number of responses in this study was 1490 elementary school teachers in fourteen areas of Bogor Regency. The modified instrument consists of 43 items to measure the TPACK ability.

Table 1: The TPACK instrument grid for elementary school teachers for math subjects

Aspect of TPACK	Indicator	Item	Total
<i>Pedagogical Knowledge (PK)</i>	Understanding of using models, strategies, learning methods to present/teach subjects	1, 2, 3, 4, 5, 6, 7, 8	8
<i>Content Knowledge (CK)</i>	Good mastery of the material with the latest and accurate references	32, 33, 34, 35, 36, 37, 38	7
<i>Technological Knowledge (TK)</i>	Knowledge of how to design learning media, using MS word to document learning administration (designing lesson plans and syllabus), understanding accessing the internet, displaying teaching materials in powerpoint (PPT) form	16, 17, 18, 19, 20, 21, 22, 23, 24	9
<i>Pedagogical Content Knowledge (PCK)</i>	Knowledge of the use of analogies in teaching and providing concrete examples in everyday life so that the material is easy to understand.	25, 26, 27, 28	4
<i>Technological Pedagogical Knowledge (TPK)</i>	Student learning by using ICT as a cognitive tool, and an assistant in finding references to create learning tools. ICT also supports collaborative learning.	29, 30, 31	3
<i>Technological Content Knowledge (TCK)</i>	Knowledge of the selection of suitable media to be used based on the material being studied, for example, learning videos are used to explain abstract material	9, 10, 11, 12, 13, 14, 15	7
<i>Technological Pedagogical Content Knowledge (TPACK)</i>	Knowledge of the use of various learning media in facilitating teachers in teaching a material.	39, 40, 41, 42, 43	5
Total			43

In this study, the data were analyzed by Rasch modeling using Winsteps software version 4.0.1. Rasch modeling was chosen as data processing because it can overcome the basic difficulty of quantitatively weighting latent qualitative phenomena (Wibisono, 2019). The advantages of the Rasch model can meet the five principles of the measurement model, namely: (1) can provide a linear scale with equal intervals, (2) predict the missing data, (3) has an estimation accuracy. (4) detect model precision, and (5) produce a standard error measurement value for the instrument used so that the calculation accuracy occurs (Sumintono, 2016).

Finding

The analysis was carried out with data sourced from 1490 elementary school teachers and 43 items. Data tabulated in Ms. software. Excel to then be converted and analyzed with the help of Winstep

4.0.1 software. Rasch modelling data processing with the help of Winstep to find out; unidimensionality, monotonic nature, item fit and person fit on Wright's map, statistical summary analysis table, and test information function analysis on the TPACK instrument for elementary school teachers for Mathematics.

One of the advantages of Rasch modeling is unidimensional measurements. Unidimensional is used to find out whether the instrument measures what it is supposed to measure (Propagation, 2018). In this study, the instrument aims to measure the TPACK of elementary school teachers for Mathematics subjects. Matlock and Turner (2016) states unidimensionality as a measuring dimension that only measures one characteristic among test takers (Sumintono, 2016). Unidimensional is the core determinant of parameter estimation (Sinnema et al., 2017). Evidence of internal consistency can be determined by unidimensional (Huberty et al., 2013). The fulfillment of the construct validity test that can measure the latent construct is a basic requirement of an instrument. In Rasch modeling, these conditions are met for unidimensional verification and local independence. The results of unidimensional calculations on a scale of 1 to 5 are shown in table 2 below:

Table 2: Unidimensional scale 1 to 5

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = PERSON information units				
	Eigenvalue	Observed	Expected	
Total raw variance in observations =	1478.2629	100.0%	100.0%	
Raw variance explained by measures =	573.2630	38.8%	38.7%	
Raw variance explained by persons =	192.6071	13.0%	13.0%	
Raw Variance explained by items =	380.6558	25.8%	25.7%	
Raw unexplained variance (total) =	905.0000	61.2%	100.0%	61.3%
Unexplned variance in 1st contrast =	63.0978	4.3%	7.0%	
Unexplned variance in 2nd contrast =	56.2803	3.8%	6.2%	
Unexplned variance in 3rd contrast =	51.5544	3.5%	5.7%	
Unexplned variance in 4th contrast =	45.3283	3.1%	5.0%	
Unexplned variance in 5th contrast =	44.2496	3.0%	4.9%	

Unidimensionality can be seen in the "raw variance explained by measure" which is located in the "observed" column. The results in table 2 show the unidimensionality of 38.8%. The requirements for the instrument to meet unidimensionality have a value greater than 20% (Shih et al., 2013). This means that the unidimensional value of the TPACK instrument with a Likert scale of five categories meets the requirements. Furthermore, to prove construct validity, it can be proven through dimensional analysis. The Eigenvalue units column is used to prove the dimensional analysis (Huberty et al., 2013). In table 2, the eigenvalue units are sequentially: 63.1, 56.3, 51.56, 45.33, and 44.2. Another proof for unidimensionality is to look at the unexplained variance (Sinnema et al., 2017) revealed that the criteria for the Eigenvalue units in the observed column are less than 15%. Table 2 shows the variance values that cannot be explained sequentially, namely: 4.3%, 3.8%, 3.5%, 3.1%, and 3.0%. The value of variance is very strong if it is in the range of 3-5% (Seol, 2016). It can be concluded that the TPACK instrument for elementary school teachers for science subjects empirically fulfills construct validity.

The use of a Likert scale in this research instrument needs to be verified so as not to confuse the respondents. For this reason, it is necessary to analyze the validity of the rating scale. Verification of the rating scale on the instrument used can be analyzed using Rasch modeling (Propagation, 2018). The monotonic nature of Rasch Modeling aims to analyze whether respondents can distinguish alternative answer choices and verify the level of agreeing response based on a five-point scale (Sulaeman, n.d. 2020). Testing the measurement rating scale on the Winsteps program, using Rating (partial credit), results are shown on the average value (observed average) and Andrich Threshold (Sumintono, 2014). Linacre (2011) confirms the Rasch-Andrich threshold to determine the size of the rating validity by showing the transition of respondents in making decisions.

The monotonic nature of the Rasch modeling of the development of the TPACK instrument using a five-category Likert scale is shown in table 3 below:

Table 3: Five-category Likert Scale

Alternative Answer Options				
1	2	3	4	5
Strongly disagree	disagree	Just agree	Agree	Strongly agree

Testing the emotional condition of respondents in verifying answer preferences with a five-category Likert scale can be seen in the monotonic nature table through the Andrich Threshold below:

Table 4: Monotonism via Andrich Threshold

SUMMARY OF CATEGORY STRUCTURE. Model="R"

CATEGORY LABEL	SCORE	OBSERVED COUNT	OBSVD %	SAMPLE AVRG	EXPECT	INFINIT MNSQ	OUTFIT MNSQ	ANDRICH THRESHOLD	CATEGORY MEASURE
1	1	1020	3	-.78	-.74	.98	.99	NONE	(-3.48)
2	2	6470	19	-.11	-.15	1.04	1.05	-2.28	-1.43
3	3	10472	30	.37	.41	.95	.94	-.35	.09
4	4	11144	32	.98	.96	.97	.98	.62	1.46
5	5	5321	15	1.59	1.59	1.01	1.02	2.01	(3.27)
MISSING		1	0	.38					

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

Winsteps analysis shows that the Andrich Threshold value moves monotonically from NONE to the negative logit direction (-2.28) and leads to the positive logit (2.01). The Observed Average column shows an increase in value from a large negative value to a large positive value. The logit value starts at -0.78 for the choice of category one (very bad), -0.11 for the choice of category two (not good), 0.37 for the choice of category three (good enough), 0.98 for the choice of category four (good), and 1.59 for the fifth choice category (very good). The increase in the logit value indicates that respondents can distinguish alternative answer choices and verify respondents who agree based on a five-category scale (Sulaeman, n.d.).

Another advantage of Rasch modeling in addition to providing unidimensional and monotonic measurements is that it can see the quality of the item fit with the model (item fit), describes whether the item is functioning normally performs a measurement. The reference for seeing fit items is by looking at the Outfit Mean Square (MNSQ) value, Outfit Z Standard Estimate (ZSTD), and Point Measure Correlation. (DiStefano & Morgan, 2010). These three references can be used to check the mismatch index. MNSQ can be accepted or declared fit when it is at a value of 0.5 logit-1.5 logit (Seol, 2016). ZSTD is acceptable if it is in the value of -1.96 logit to +1.96 logit (Gomez et al., 2012). Point Measure Correlation criteria can be accepted in the range of 0.32 logit < x < 0.8 logit (Boone & Noltemeyer, 2017). This study tested 43 items on the TPACK instrument for elementary school teachers for Mathematics. The results of the analysis using the Winsteps version 4.0.1 program obtained 38 items that were declared fit and 5 items that were not fit are presented in table 5 below:

Table 5: Items fit and items do not fit five categories Likert scale

Nilai outfit MNSQ	Nilai PT-Measure Correlation	Misfit	Item fit
0,68 logit to 1,48 logit	0,50 logit to 0,64 logit	B5, B9,B10, B11, B27	B1,B2,B3,B4,B6,B7,B8,B12,B13, B14,B15,B16,B17,B18,B19,B20, B21,B22,B23,B24,B25,B26,B28, B29,B30,B31,B32,B33,B34,B35, B36,B37,B38,B39,B40,B41,B42, B43
Total		5	38

Based on table 5, information is obtained regarding the suitability of items with the model, namely there are 38 items that fit the model on a five-category Likert scale. There are 5 items that do not match the model, namely: B5, B9, B10, B11, and B27. The MNSQ values sequentially start from 0.68 logit to 1.48 logit and the PT-Measure Correlation value is in the range of 0.50 logit to 0.64 logit. It was decided that five items that did not fit were dropped out or not used.

Through Rasch modelling, it can be seen that teachers have inappropriate response patterns. The pattern of inappropriate responses is the discrepancy of the answers given based on their abilities compared to the ideal model (Sumintono, 2014). The pattern of inappropriate responses is the discrepancy of the answers given based on their abilities compared to the ideal model (Sulaeman, n.d.). To detect unsuitable teachers or non-fit persons by looking at the MNSQ, ZSTD, and Point Measure Correlation values (Boone & Noltemeyer, 2017). The value criteria used are: for MNSQ the accepted value is $0.5 < \text{MNSQ} < 1.5$, the ZSTD value is $-2.0 < \text{ZSTD} < +2.0$, and for the PT-Measure Correlation value $0.4 < \text{Pt Measure Corr} < 0.85$ (Bambang Sumintono & Wahyu Widhiarso, 2015). Note that the sample size can affect the ZSTD value. Large samples will produce ZSTD values above 3, some experts recommend not using the ZSTD value if the sample size to be calibrated is $N > 500$ (Bambang Sumintono & Wahyu Widhiarso, 2015). In this study, to determine person fit, the researcher used the MNSQ outfit outfit criteria. The results of the analysis with winsteps obtained the data of person fit and not fit as follows:

Table 6: Person fit and Person not fit five categories Likert scale

Nilai outfit MNSQ (Person fit)	Nilai PT-Measure Correlation (Person fit)	Person Misfit	Person fit
0,55 logit to 1,42 logit	0,46 logit to 0,60 logit	584	906

Based on table 6, it is obtained that there are 584 people who do not fit and who are fit 906 responses. The Outfit NNSQ value sequentially starts from 0.55 logit to 1.42 logit and the PT-Measure Correlation value is in the range 0.46 logit to 0.60 logit. It can be concluded that the number of responses that dropped out was 584 responses.

Proof of person analysis can also be done using a person map. The advantage of the person map is that it can detail each respondent by sorting high and low ability respondents. Based on Winstep analysis on Output: person (fit) order and Output: person measure obtained by the category of respondents' ability grouping as follows:

Table 7: The TPACK level for Mathematics is based on the response ability of primary school teachers

Mean	= 0,55				
SD	= 0,72				
Respon	= 906				
Respondent's ability level	Logit calculation	Skala Logit Measure	Total	Percentage	
Very high	1,99 logit + 0,72 logit	2,71	48	5,30	
High	1,27 logit + 0,72 logit	1,99	96	10,60	
Currently	0,55 logit + 0,72 logit	1,27	540	59,60	
Low	0,55 logit - 0,72 logit	-0,17	190	20,97	
Very low	-0,17 logit - 0,72 logit	-0,89	32	3,53	
Total			906	100	

Table 7 shows the categories of teachers' abilities in answering the items of the TPACK instrument with a Likert scale of five response categories. From 906 teachers who became respondents information was obtained; (1) teachers with very high TPACK ability of 5.30%, This means that the frequency of the teacher's ability level is very high in determining the choice of the right response category in each statement item, (2) A total of 10.60% of teachers have high ability in determining the

choice of the right response category in each statement item, (3) 59.60% of teachers have moderate ability in determining the choice of the right response category in each statement item, (4) The number of teachers with low ability in determining the choice of the right response category in each statement item is 20.97%, and (5) with a very low ability in determining the choice of the right response category in each statement item, the percentage is 20.97%. Based on these data, teachers with moderate ability.

In the Rasch modeling, information that describes the distribution of teacher abilities and the distribution of item difficulty levels with the same scale can be seen in the Output tables: 12 item map. Karami, (2015) states that Rasch modeling through a variable map plot provides an overview of persons and items in the same logit interval scale. The output of Winsteps version 4.01 shows fit and person fit items on the Wright map as follows:

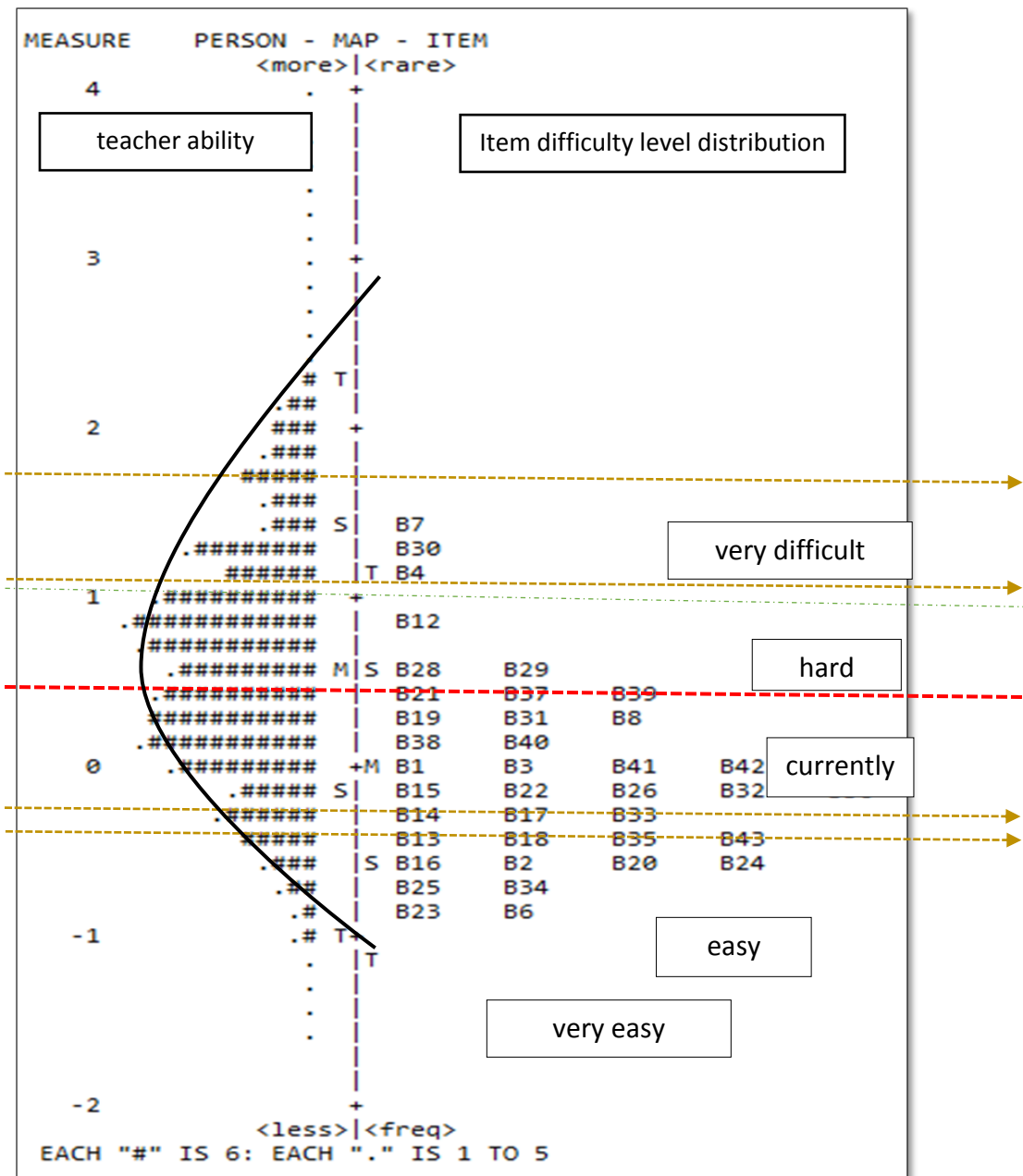


Figure 2: Item-person variable map Likert scale five response categories

Figure 2 is the output of Winsteps version 4.0.1 which presents the test results of 38 fit items and 906 person fit. The left side is the distribution of the ability of teachers (respondents), while the right side of the distribution of items. Grouping the ability of items in categories: very difficult, difficult, moderate, easy, and very easy to calculate based on the statement item analysis on the item measure. Mean and Standard Deviation (SD) when combined, the level of statement items can be grouped (Sulaeman, n.d., 2020). Boone & Noltemeyer (2017) express zero logit is location “M” or item mean. The difficulty grouping of items is presented in table 8 below:

Table 8: Item difficulty grouping

Mean item measure	= 0,0 logit		
Standar Deviasi (SD)	= 0,56 logit		
Item group limit		Item difficulty level	Item Number
very difficult	1,12 logit + 0,56 logit	1,68 logit	B7,B30
Hard	0,56 logit + 0,56 logit	1,12 logit	B4,B12
Currently	0,0 logit + 0,56 logit	0,56 logit	B28, B29,B21,B37,B39,B19,B31,B8,B38,B 40,B1,B3,B41,B42,B15,B22,B26,B32, B36,B14,B17,B33,B43,B13,B18,B35, B16,B2,B20,B24
Easy	0,0 logit - 0,56 logit	-0,56 logit	B25, B34
Very easy	-0,56 logit - 0,56 logit	-1,12 logit	B23, B6
Total			38

Table 8 provides information on items B7 and B30 which are very difficult items for teachers to respond to in determining the choice of response categories. Items B23 and B6 are items that are very easy for the teacher to respond to in determining the choice of item categories. In general, the item-person variable map concluded that the teacher's ability was higher than the level of difficulty of the items in the questionnaire given.

Rasch modeling with Winsteps software in this study provides information on the results of measuring the validity and reliability of the TPACK ability of elementary school teachers in Mathematics through statistical summary outputs. Overall, the summary statistics provide information about the quality of the teacher's response patterns as respondents, the quality of the instrument used, and the interaction between the person and the item (Bambang Sumintono & Wahyu Widhiarso, 2015). Internal reliability information can be obtained from the output summary statistics (Sulaeman, n.d. 2020). The data results from 906 person fit and 38 fit items on the TPACK instrument are presented in table 9 below:

Table 9: Summary Statistics Likert scale five response categories

Table 9 displays two summary statistics tables from Rasch MIRT software. The top table is for 906 measured persons, and the bottom table is for 38 measured items. Red boxes highlight key values, and red arrows point to labels: 'person measure' (MEASURE), 'person separation' (SEPARATION), 'person reliability' (PERSON RELIABILITY), 'Cronbach Alpha' (CRONBACH ALPHA), 'item calibration' (MEASURE), 'SE=standar error' (S.E.), and 'item reliability' (ITEM RELIABILITY).

SUMMARY OF 906 MEASURED (EXTREME AND NON-EXTREME) PERSON								
	TOTAL SCORE	COUNT	MEASURE	MODEL S. E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	128.7	38.0	.64	.20				
P. SD	20.2	.0	.82	.06				
S. SD	20.2	.0	.83	.06				
MAX.	190.0	38.0	7.02	1.83				
MIN.	75.0	37.0	-1.53	.19				
REAL RMSE	.22	TRUE SD	.79	SEPARATION	3.59	PERSON RELIABILITY	.93	
MODEL RMSE	.21	TRUE SD	.80	SEPARATION	3.81	PERSON RELIABILITY	.94	
S. E. OF PERSON MEAN = .03								
PERSON RAW SCORE-TO-MEASURE CORRELATION = .98								
CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .94 SEM = 5.13								
SUMMARY OF 38 MEASURED (NON-EXTREME) ITEM								
	TOTAL SCORE	COUNT	MEASURE	MODEL S. E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	3067.3	906.0	.00	.04	1.00	-.3	1.00	-.2
P. SD	344.6	.2	.56	.00	.18	4.0	.18	4.0
S. SD	349.2	.2	.56	.00	.18	4.1	.19	4.1
MAX.	3599.0	906.0	1.40	.04	1.49	9.5	1.48	9.3
MIN.	2206.0	905.0	-.89	.04	.68	-8.4	.68	-8.4
REAL RMSE	.04	TRUE SD	.55	SEPARATION	13.24	ITEM RELIABILITY	.99	
MODEL RMSE	.04	TRUE SD	.55	SEPARATION	13.69	ITEM RELIABILITY	.99	
S. E. OF ITEM MEAN = .09								

Table 9 provides four important information, namely: (1) Person reliability and item reliability, (2) Person and item separation index, (3). Precision of measurement, and (4) Person Measure.

Person reliability and item reliability, according Maat & Rosli (2016) the reliability value is considered internally consistent if it is close to one. The stability of persons and items on Rasch's reliability is in the range of zero to one, which is interpreted as Cronbach's Alpha (Boone & Noltemeyer, 2017). From table 9, the person reliability index value is 0.93, the item reliability index is 0.99 and the Cronbach Alpha coefficient is 0.94. It can be concluded that the high reliability estimate means that there is a consistent interaction between the teacher's response and the item. Other evidence regarding the reliability of persons and items refers to the following criteria for the value of person reliability and item reliability:

Table 10: Criteria for the value of person reliability and item reliability

Indeks	Category
< 0,67	weak
0,67 - 0,80	enough
0,80 - 0,90	good
0,91 - 0,94	very good
> 0,94	special

Sumber, (Bambang Sumintono & Wahyu Widhiarso, 2015)

Based on the criteria for the value of person reliability and item reliability, the criteria are very good, Thus the TPACK instrument for elementary school teachers with subject matter is considered reliable to use.

Person and item separation index, provide an estimate of the teacher's ability to answer the items of the TPACK questionnaire instrument. Perera et al., (2018) states that the probability of the spread of respondents responding to items correctly and how wide the spread of items from easy to difficult items is determined from the greater the person index and item separation. According Duncan et al., (2003) index criteria 1.50 acceptable, the range of 2.00 is good, and the index meets the criteria of

3.00 very good. Based on table 9 the index of person separation is 3.56 and the index of item separation is 13.24, it can be said that the instrument is reliable to measure the TPACK of elementary school teachers.

Precision of measurement describe a strong trust in the instrument to draw conclusions or decisions (Sulaeman, n.d. 2020). To evaluate the reliability and strength of the instrument, accurate and reliable measurements are needed (Perera et al., 2018), Perera further emphasized that an instrument must have a standard error of less than 0.5 (Perera et al., 2018). Based on table 9 in the column "SE Model" it is 0.04 which means that the TPACK instrument is reliable and has good measurements.

Person measure used to see the tendency of respondents to answer agree. Bambang Sumintono & Wahyu Widhiarso (2015) states that the average value of the person measure is more than 0.0 logit, indicating the tendency of respondents to agree on the statement item. Based on table 9, the value is + 0.64 logit which means that the teacher (respondent) tends to answer agreeably on the TPACK instrument.

Rasch modeling with Winsteps software can provide information about the measurement results in the form of a Test Information Function graph. Bambang Sumintono & Wahyu Widhiarso (2015) confirms that the Test Information Function graph provides measurement focus information, the function of the measurement (initial screening, remedial, or test) that the researcher performs, and the reliability of the measurement performed. Based on data from 1490 respondents and 43 items, the output of the TPACK instrument information function with a Likert scale of five response categories is obtained as follows:

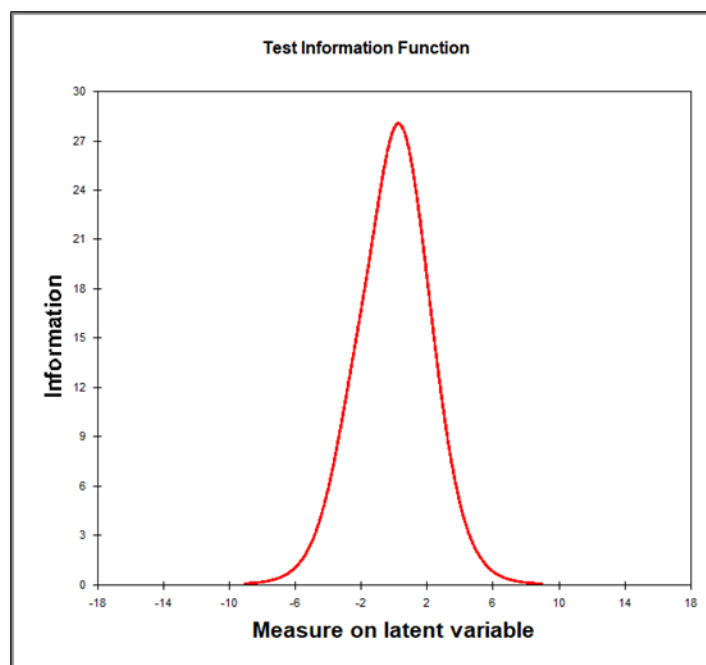


Figure 3; Information function Likert scale five response categories

Figure 3 shows a high peak, the peak height in the information function describes the amount of information provided (Bambang Sumintono & Wahyu Widhiarso, 2015). The higher the peak of the information, the higher the information provided (Wahyuningsih, 2020). The graph of the information function of the Likert scale for five response categories provides measurement information obtained from the TPACK instrument for elementary school teachers for Mathematics. The X axis shows the teacher's ability to work on the TPACK questionnaire given. The y-axis shows the value and function of the information. Based on the graph, the TPACK ability level of elementary school teachers in Mathematics

subjects is at the medium ability level, based on the graph, the TPACK ability level of elementary school teachers in Mathematics subjects is at the medium ability level.

Conclusion

The questionnaire instrument used to measure the TPACK ability of elementary school teachers in Mathematics is fit with the Rasch model to measure the TPACK ability of teachers with moderate ability. This is evidenced by the results of the analysis of the person measure index data of +0.64 logit, which means that it shows the tendency of respondents to agree on the statement item. The TPACK instrument used has a good level of reliability and measurement, this is evidenced by the standard error index (SE) of only 0.04, the person separation index is 3.56 and the item separation index is 13.24. The instrument used in this study has high reliability, data analysis shows a person reliability index of 0.93, an item reliability index of 0.99 and a Cronbach Alpha coefficient of 0.94. These three estimates show a high reliability score or mean that there is a consistent interaction between the teacher's response and the item. Thus the instrument serves to measure the TPACK ability of elementary school teachers in Mathematics.

From this research, it is hoped that many similar studies will develop instruments to measure the ability of teachers, especially elementary school teachers. Measurement of teacher abilities can certainly be used as a reference in improving the quality of education and provide an overview for the government in determining teacher professionalism development policies.

Acknowledgement

The Authors would like to thank you to Head of Department of Education Bogor District (Juanda Dimansyah, S.E., M.M.) and Head of BKPSDM Bogor District (Drs. Zulkifli, A.P., M.M., M.Si.). Who has provided the opportunity and support to the author to continue his education to the master's level, as well as the study scholarships provided.

References

- Akbulut, Y. (2007). Implications of two well-known models for instructional designers in distance education: Dick-carey versus Morrison-ross-kemp. *Turkish Online Journal of Distance Education*, 8(2), 62–68. <https://doi.org/10.17718/tojde.13470>
- Ariani, D. N. (2015). Hubungan antara Technological Pedagogical Content Knowledge dengan Technology Integration Self Efficacy Guru Matematika di Sekolah Dasar. *Muallimuna*, 1(1), 79–91. <https://doi.org/10.31602/muallimuna.v1i1.277>
- Badan Standar Nasional Pendidikan. (2010). Paradigma Pendidikan Nasional Abad XXI. *Paradigma Pendidikan Nasional Abad XXI*, 1–59.
- Bambang Sumintono & Wahyu Widhiarso. (2015). *Aplikasi Pemodelan Rasch Pada Assesment Pendidikan (Pertama)*. Trim Komunikata.
- Boone, W. J., & Noltemeyer, A. (2017). Rasch analysis: A primer for school psychology researchers and practitioners. *Cogent Education*, 4(1), 1–13. <https://doi.org/10.1080/2331186X.2017.1416898>
- Caturiyati, K. H. dan. (2013). Validitas Konstruk (construct validity) dalam Pengembangan Instrumen Penilaian Non-Kognitif. *Journal of Chemical Information and Modeling*, 53(9), 1689–1699.
- Das, S. W. H., Halik, A., Nasir, M., & Suredah, S. (2009). Pencapaian Kompetensi Guru Sekolah Dasar Negeri. *Seminar Nasional Pendidikan, Sains dan Teknologi Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Muhammadiyah Semarang*, 350–359.
- Desstya, A. (2018). Validitas Reliabilitas Instrument Technological Pedagogical Content Knowledge (Tpack) Guru Sekolah Dasar Muatan Pelajaran Ipa. *AL-ASASIYYA: Journal Of Basic Education*, 3(1), 126. <https://doi.org/10.24269/ajbe.v3i1.1338>
- DiStefano, C., Greer, F. W., Kamphaus, R. W., & Brown, W. H. (2014). Using Rasch Rating Scale

- Methodology to Examine a Behavioral Screener for Preschoolers At Risk. *Journal of Early Intervention*, 36(3), 192–211. <https://doi.org/10.1177/1053815115573078>
- DiStefano, C., & Morgan, G. B. (2010). Evaluation of the BESS TRS-CA Using the Rasch Rating Scale Model. *School Psychology Quarterly*, 25(4), 202–212. <https://doi.org/10.1037/a0021509>
- Duncan, P. W., Bode, R. K., Lai, S. M., & Perera, S. (2003). Rasch analysis of a new stroke-specific outcome scale: The stroke impact scale. *Archives of Physical Medicine and Rehabilitation*, 84(7), 950–963. [https://doi.org/10.1016/S0003-9993\(03\)00035-2](https://doi.org/10.1016/S0003-9993(03)00035-2)
- Fusarelli, L. D. (2008). *Innovation, NCLB, and the Fear Factor*.
- Gomez, L. E., Arias, B., Verdugo, M. Á., & Navas, P. (2012). An Outcomes-Based Assessment of Quality of Life in Social Services. *Social Indicators Research*, 106(1), 81–93. <https://doi.org/10.1007/s11205-011-9794-9>
- Han, E. S., & goleman, daniel; boyatzis, Richard; Mckee, A. (2019). 濟無No Title No Title. In *Journal of Chemical Information and Modeling* (Vol. 53, Nomor 9).
- Harris, J., Koehler, M., & Koehler, M. J. (2009). *Apa Itu Konten Pedagogis Teknologi*. 9, 60–70.
- Huberty, J., Vener, J., Gao, Y., Matthews, J. L., Ransdell, L., & Elavsky, S. (2013). Developing an instrument to measure physical activity related self-worth in women: Rasch analysis of the Women's Physical Activity Self-Worth Inventory (WPASWI). *Psychology of Sport and Exercise*, 14(1), 111–121. <https://doi.org/10.1016/j.psychsport.2012.07.009>
- Hy, E. D. W., Orks, I. T. W., & Tidak, A. (2015). *Teknologi Digital dalam Pendidikan Matematika: Mengapa Berfungsi (Atau Tidak)*.
- Jang, S. J., & Tsai, M. F. (2012). Exploring the TPACK of Taiwanese elementary mathematics and science teachers with respect to use of interactive whiteboards. *Computers and Education*, 59(2), 327–338. <https://doi.org/10.1016/j.compedu.2012.02.003>
- Karami, H. (2015). Book review: Rasch Analysis in the Human Sciences. *Language Testing*, 32(4), 545–548. <https://doi.org/10.1177/0265532214567642>
- Koehler, M. J. (2006). *Konten Pedagogis Teknologi Pengetahuan : Sebuah Kerangka Pengetahuan Guru*. 108, 1017–1054.
- Koehler, M. J., Michigan, U. N., Mishra, P., & Arizona, U. N. (2017). *Apa Yang Terjadi Saat Guru Merancang Teknologi Pendidikan ? Perkembangan Pengetahuan Konten Pedagogis Teknologi*.
- Koh, J. H. L., Chai, C. S., & Lee, M. H. (2015). Technological Pedagogical Content Knowledge (TPACK) for Pedagogical Improvement: Editorial for Special Issue on TPACK. *Asia-Pacific Education Researcher*, 24(3), 459–462. <https://doi.org/10.1007/s40299-015-0241-6>
- Ling Koh, J. H., Chai, C. S., & Tay, L. Y. (2014). TPACK-in-Action: Unpacking the contextual influences of teachers' construction of technological pedagogical content knowledge (TPACK). *Computers and Education*, 78, 20–29. <https://doi.org/10.1016/j.compedu.2014.04.022>
- Maat, S. M., & Rosli, M. K. (2016). The Rasch Model Analysis for Statistical Anxiety Rating Scale (STARS). *Creative Education*, 07(18), 2820–2828. <https://doi.org/10.4236/ce.2016.718261>
- Nasional, S., Guru, P. K., Ukg, M., Ptbb, J., & Uny, F. T. (2012). *Seminar Nasional 2012 "Peningkatan Kompetensi Guru dalam Menghadapi UKG" Jurusan PTBB FT UNY, 15 Desember 2012 I*. 1–11.
- Niess, M. L., Ronau, R. N., Shafer, K. G., Driskell, S. O., Harper, S. R., Johnston, C., & Barat, U. M. (2009). *Standar TPACK Guru Matematika dan Model Pengembangan*. 9, 4–24.
- Perera, H. N., Granziera, H., & McIlveen, P. (2018). Profiles of teacher personality and relations with teacher self-efficacy, work engagement, and job satisfaction. *Personality and Individual Differences*, 120(September 2017), 171–178. <https://doi.org/10.1016/j.paid.2017.08.034>
- Pilihan, D. A. N., Untuk, G., & Learning, M. S. (2015). Perbandingan Akurasi Penggunaan Skala Likert Dan Pilihan Ganda Untuk Mengukur Self-Regulated Learning. *Jurnal Kependidikan: Penelitian Inovasi Pembelajaran*, 45(2), 118670. <https://doi.org/10.21831/jk.v45i2.7493>
- Propagation, W. (2018). *M 2 , P 2 M 1 , P 1*. May, 2–3.
- Putrawangsa, S., & Hasanah, U. (2018). Integrasi Teknologi Digital Dalam Pembelajaran Di Era Industri 4.0 Kajian dari Perspektif Pembelajaran Matematika. *Jurnal Pemikiran dan Penelitian Pendidikan*, 16(1), 42–54.
- Richardo, R. (2016). Program Guru Pembelajar: Upaya Peningkatan Guru Profesionalisme Guru Abad

21. *Prosiding Seminar Matematika dan Pendidikan Matematika, November*, 777–785.
- Sari, A. P. P., Amin, M., & Lukiati, B. (2017). Buku Ajar Bioteknologi Berbasis Bioinformatika Dengan Model ADDIE. *Jurnal Pendidikan: Teori, Penelitian, dan Pengembangan*, 2(6), 769.
- Seol, H. (2016). Using the bootstrap method to evaluate the critical range of misfit for polytomous rasch fit statistics. *Psychological Reports*, 118(3), 937–956. <https://doi.org/10.1177/0033294116649434>
- Shih, C. L., Chen, C. H., Sheu, C. F., Lang, H. C., & Hsieh, C. L. (2013). Validating and improving the reliability of the EORTC QLQ-C30 using a multidimensional rasch model. *Value in Health*, 16(5), 848–854. <https://doi.org/10.1016/j.jval.2013.05.004>
- Sinnema, C., Meyer, F., & Aitken, G. (2017). Capturing the Complex, Situated, and Active Nature of Teaching Through Inquiry-Oriented Standards for Teaching. *Journal of Teacher Education*, 68(1), 9–27. <https://doi.org/10.1177/0022487116668017>
- Sintawati, M., & Indriani, F. (2019). Pentingnya Literasi ICT Guru di Era Revolusi Industri 4.0. *Jurnal Ilmiah Pendidikan Matematika*, 1(2), 417–422.
- Sulaeman, E. (n.d.). *THE FUNCTIONALITY OF THE MIDDLE VALUE OF THE INDONESIAN VERSION OF EMOTIONAL LEARNING INSTRUMENT*. 6(1), 26–46.
- Sumintono, B. (2014). Model Rasch untuk Penelitian Sosial Kuantitatif. *Makalah kuliah umum di jurusan statistika, ITS Surabaya, 21 November 2014, November 201*, 1–9. <http://deceng3.wordpress.com>
- Sumintono, B. (2016). Aplikasi pemodelan Rasch pada asesmen pendidikan: Implementasi penilaian formatif (Assessment for learning). *Makalah dipresentasikan dalam Kuliah Umum pada Jurusan Statistika, Institut Teknologi Sepuluh November, Surabaya, 17 Maret 2016., March*, 1–19. http://eprints.um.edu.my/15876/1/ITS_rasch_model_asesment_for_learning.pdf
- Triyono. (2003). Teknik Sampling Dalam Pelaksanaan Penelitian. *Info Kesehatan*, 7(1), 64. <https://osf.io/preprints/inarxiv/dcq8u/download>
- Wahyuningsih, S. A. (2020). Penggunaan Model Rasch Untuk Analisis Instrumen the Use of Rasch Model for Analyzing Test. *Jurnal Pendidikan Matematika*, 3(1), 45–50.
- Wibisono, S. (2019). Aplikasi Model Rasch Untuk Validasi Instrumen Pengukuran Fundamentalisme Agama Bagi Responden Muslim. *Jurnal Pengukuran Psikologi dan Pendidikan Indonesia (JP3I)*, 3(3). <https://doi.org/10.15408/jp3i.v3i3.10731>
- Xu, M., David, J. M., & Kim, S. H. (2018). *Keempat Revolusi Industri : Peluang dan Tantangan*. 9(2), 57–62.
- Yu, C. H. (2012). *A simple guide to Item Response Theory (IRT) and Rasch Modeling*. 30. <http://www.creative-wisdom.com/computer/sas/IRT.pdf>
- Zamzam, I., Mahdi, S. A., & Ansar, R. (2017). *Jurnal Ilmiah Akuntansi Peradaban 1*. 1–24.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).